

RÈGLE D'ARRÊT (G7)

(21 / 09 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) En **théorie séquentielle** (cf **problème de décision séquentielle**) on considère un **espace probabilisé** $(\Omega, \mathcal{F}, P)$, une **filtration** $\mathcal{F} = (\mathcal{F}_n)_{n \in \mathbf{N}}$ sur $\mathcal{F}$, une suite d'**espaces d'état** $(\mathcal{X}_n, \mathcal{B}_n)_{n \in \mathbf{N}}$ et un **processus stochastique** $X = (X_n)_{n \in \mathbf{N}}$, avec $X_n : \Omega \mapsto \mathcal{X}_n, \forall n \in \mathbf{N}$.

On appelle alors :

(a) **règle de décision d'arrêt**, ou simplement **règle d'arrêt**, une fonction mesurable $N : \Omega \mapsto \mathbf{N}$ (cf **application mesurable**) tq :

$$(0) \quad \begin{aligned} [N = n] &= \{\omega \in \Omega : N(\omega) = n\}, \quad \forall n \in \mathbf{N}, \\ P(N = \infty) &= 0; \end{aligned}$$

(b) **règle de décision d'arrêt mixte**, ou simplement **règle d'arrêt mixte**, une **suite** $m(x) = (m_N(x_1, \dots, x_N))_{N \in \mathbf{N}^*}$ d'**applications mesurables** :

$$(1) \quad m_N : \mathcal{X}_1 \times \dots \times \mathcal{X}_N \mapsto [0, 1],$$

dans laquelle chaque $m_N(x_1, \dots, x_N)$ est la **probabilité conditionnelle** d'arrêter l'**échantillonnage** sachant qu'on a observé N réalisations $X_n(\omega) = x_n$ de $X, \forall n \in \mathbf{N}^*$.

On note simplement m_0 la probabilité de ne pas procéder à un échantillonnage.

(ii) Une **règle d'arrêt mixte** se définit, de façon équivalente, à l'aide d'une suite d'applications $m'(x) = (m'_N(x_1, \dots, x_N))_{N \in \mathbf{N}^*}$ tq :

$$(2) \quad m'_N : \mathcal{X}_1 \times \dots \times \mathcal{X}_N \mapsto [0, 1],$$

où $m'_N(x_1, \dots, x_N)$ représente la probabilité conditionnelle de ne pas arrêter l'échantillonnage après observation des $N-1$ premières réalisations et d'arrêter après la N -ième, sachant qu'on a observé les N réalisations $X_n(\omega) = x_n, \forall n \in \mathbf{N}^*$.

(iii) On établit les **relations de « dualité »** suivantes :

$$(3) \quad \begin{aligned} m'_0 &= m_0, \\ m'_N(x_1, \dots, x_N) &= p_N(x), \end{aligned}$$

où $p_N(x) = (1 - m_0) \cdot \{\prod_{n=1}^{N-1} (1 - m_n(x_1, \dots, x_n))\} \cdot m_N(x_1, \dots, x_N)$.

Si $v : \Omega \mapsto \mathbf{N}$ est une variable (aléatoire) entière représentant la taille (aléatoire) de l'**échantillon**, la **probabilité d'arrêt de l'échantillonnage** après la N-ième observation est définie selon :

$$(4) \quad Q_\theta ([v = N]) = E_\theta m_N' (X_1, \dots, X_N), \quad \forall N \in \mathbf{N}^*, \forall \theta \in \Theta.$$

La probabilité de procéder à un **échantillonnage fini** (ie la probabilité de l'évènement $[v < +\infty]$) est alors :

$$(5) \quad Q_\theta ([v < +\infty]) = \sum_{N \in \mathbf{N}} Q_\theta ([v = N]), \quad \forall \theta \in \Theta.$$

Cette variable v est appelée **temps d'arrêt**, ou **taille d'arrêt**, ou encore **indice d'arrêt**. On a donc, par définition :

$$(6) \quad \begin{aligned} P_\theta ([v = N] / [X = x]) &= m_N' (x), & \forall (N, \theta) \in \mathbf{N} \times \Theta, \\ P_\theta ([v = +\infty] / [X = x]) &= 1 - \sum_{N \in \mathbf{N}} m_N' (x), & \forall \theta \in \Theta, \end{aligned}$$

où $X = (X_1, \dots, X_N)$ et $x = X(\omega) = (x_1, \dots, x_N)$. Les probabilités conditionnelles des membres de gauche ne dépendent pas de θ .

Compte tenu de (1), une **règle d'arrêt pure** est, en particulier, une suite $m = (m_N)_{N \in \mathbf{N}}$ d'applications :

$$(7) \quad m_N : \mathcal{X}_1 \times \dots \times \mathcal{X}_N \mapsto \{0, 1\}.$$

Elle correspond à une division de l'espace d'état (eg **espace d'observation**) $\mathcal{X} = \prod_{n \in \mathbf{N}} \mathcal{X}_n$ en deux zones :

(a) l'une parfois appelée **zone d'arrêt** (l'échantillonnage s'arrête si les observations « tombent » dans cette zone) ;

(b) l'autre appelée **zone de poursuite**, ou **zone de continuation**.